Santa Fe Institute Working Paper 16-02-005
arXiv:1602.08646

The Ambiguity of Simplicity in Quantum and Classical Simulation

Cina Aghamohammadi,* John R. Mahoney,† and James P. Crutchfield‡
Complexity Sciences Center and Physics Department,
University of California at Davis, One Shields Avenue, Davis, CA 95616
(Dated: December 29, 2016)

A system's perceived simplicity depends on whether it is represented classically or quantally. This is not so surprising, as classical and quantum physics are descriptive frameworks built on different assumptions that capture, emphasize, and express different properties and mechanisms. What is surprising is that, as we demonstrate, simplicity is ambiguous: the *relative* simplicity between two systems can *change sign* when moving between classical and quantum descriptions. Here, we examine the minimum required memory for simulation. We see that the notions of absolute physical simplicity at best form a partial, not a total, order. This suggests that appeals to principles of physical simplicity, via Ockham's Razor or to the "elegance" of competing theories, may be fundamentally subjective. Recent rapid progress in quantum computation and quantum simulation suggest that the ambiguity of simplicity will strongly impact statistical inference and, in particular, model selection.

Keywords: quantum information, information theory, stochastic process, hidden Markov model, ϵ-machine

PACS numbers: 03.67.-a 05.30.-d 89.75.Kd 89.70.+c

We are to admit no more causes of natural things than such as are both true and sufficient to explain their appearances.

Isaac Newton, 1687
Philosophiæ Naturalis Principia Mathematica,
Book III, p. 398 [1]

Introduction Beyond his theory of gravitation, development of the calculus, and pioneering work in optics, Newton engendered a critical abstract transition that has resonated down through the centuries, guiding and even accelerating science's growth: Physics began to perceive the world as one subject to concise mathematical Laws. Above, Newton suggests that these Laws are not only a correct perception but they are also *simple*. Consequently, one should abandon the Ptolemaic epicycles for Newton's elegant $F = ma$ and $F_g \propto m_1 m_2 / r^2$.

The desire for simplicity in a theory naturally leads us to consider *simplicity as a means for comparing* alternative theories. Here, we compare the parsimony of two descriptions of stochastic processes—one classical and one quantum. Classical versus quantum comparisons have, of late, captured our attention both for reasons of principle and of experiment. *Quantum supremacy* holds that quantum systems behave in ways beyond those that can be efficiently simulated by classical computers [2]. A single cold 2D Fermi gas supports coexistence of both quantum mechanical states at its core and classical states on its periphery [3, 4]. The overriding impression is that now is an interesting time for the foundations of quantum mechanics. The following adds a new phenomenon to these debates on the balance of classical and quantum theories, as concerns the simplicity of their descriptions.

To start, we consider a Nature full of stationary stochastic processes. A theory, then, is a mathematical object capable of yielding a process' probabilities. We can straightforwardly say that one process is more random than another via comparing their temperatures or thermodynamic entropies. But how to compare them in terms of their structural simplicities? We make use of a well developed measure of simplicity in stochastic processes—the statistical complexity—a measure of internal memory [5] or the minimum required memory to simulate a process. It provides a concrete and interpretable answer to the question, which process is structurally simpler? By applying this comparison, we may order all processes from the simplest to the most complicated [6].

With recent progress in quantum computation [7–9], an interesting twist comes about if we add quantum mechanics to our modeling toolbox. Descriptions that act on a quantum substrate offer new and surprising options. For example, it was shown that a quantum mechanical description can lead to a simpler representation [10–14] and even in some cases infinitely simpler [15, 16]. Recently, this quantum advantage was verified experimentally [17]. Proceeding with these methods, we discover what is most surprising: the *relative simplicity* of classical and quantum descriptions can change. Specifically, there are stochastic processes, A and B, for which classical theory says A is simpler than B, but quantum mechanics says B is simpler than A. What started out as a neat classical array is upended by a new quantum simplicity order. This means quantizing a simple classical model may not be as simple as quantizing a more complicated classical model. As a consequence model selection is complicated by the addition of a quantum model class.

Classical and Quantum Simplicity We consider stationary, ergodic processes: each a bi-infinite sequence of random variables $X_{-\infty:\infty} = \ldots X_{-2} X_{-1} X_0 X_1 X_2 \ldots$ where each random variable X_t takes some value x_t in a discrete alphabet set $\mathcal{A}$ and where all probabilities $\Pr(X_t, \ldots, X_{t+L})$ are time-invariant.

How is their degree of randomness quantified? Information theory [18] measures the uncertainty in a single observation X_0 via the *Shannon entropy*: $H[X_0] = -\sum_{x\in\mathcal{A}} \Pr(x)\log_2 \Pr(x)$ and the irreducible uncertainty per observation via the *entropy rate* [19]: $h_\mu = \lim_{L\to\infty} H[X_{0:L}]/L$. If we interpret the left half $X_{-\infty:0} = \ldots X_{-2}X_{-1}$ as the "past" and the right half $X_{0:\infty} = X_0X_1X_2\ldots$ as the "future", we see that the entropy rate is the average uncertainty in the next observable given the entire past: $h_\mu = H[X_0|X_{-\infty:0}]$. Thus, as we take into account past correlations, the naive uncertainty $H[X_0]$ reduces to h_μ.

How reducible is our uncertainty in the future $X_{0:\infty}$ knowing the past $X_{-\infty:0}$? The answer is given by the mutual information between the past and the future—the *excess entropy* [20]: $\mathbf{E} = \mathrm{I}[X_{-\infty:0} : X_{0:\infty}]$. With h_μ and $\mathbf{E}$, we measure randomness and predictability, respectively.

Let's say we want to simulate a given process. To do this we write a computer code that follows an algorithm and allocate the memory the algorithm needs. For a given process *computational mechanics* [21] identifies the optimal algorithm—the process' *ϵ-machine*. This is a *unifilar hidden Markov model* [22] that uses only the minimum required memory for simulation. We view a process' ϵ-machine as the "theory" of a process in that it specifies a mechanism that exactly simulates a process' behaviors. In this way, computational mechanics supplements $\mathbf{E}$ and h_μ with a measure of structure—the minimum required amount memory to simulate the given process.

The ϵ-machine consists of *causal states* $\sigma \in \boldsymbol{\mathcal{S}}$ defined by an equivalence relation $\sim$ that groups histories, say $x_{-\infty:t}$ and $x_{-\infty:t'}$, that lead to the same future predictions $\Pr(X_{t:\infty}|\cdot)$: $x_{-\infty:t} \sim x_{-\infty:t'} \iff \Pr(X_{t:\infty}|x_{-\infty:t}) = \Pr(X_{t':\infty}|x_{-\infty:t'})$. From this, one concludes that a process' ϵ-machine is, in a well defined sense, its simplest predictive theory.

Translating this notion of simplicity into a measurable quantity, we ask: *What is the minimum memory necessary to* implement *optimal prediction?* The answer is the historical information stored in the ϵ-machine. Quantitatively, this is the Shannon entropy of the causal-state stationary distribution $\{\pi_\sigma\}$, the *statistical complexity*:

$$C_\mu = \mathrm{H}[\boldsymbol{\mathcal{S}}] = -\sum_{\sigma\in\boldsymbol{\mathcal{S}}} \pi_\sigma \log_2 \pi_\sigma \ , \qquad (1)$$

It is well known that the excess entropy is a lower-bound on this structural measure: $\mathbf{E} \le C_\mu$. In fact, this relation is only rarely an equality [23]. And so, while $\mathbf{E}$ quantifies the amount to which a process is subject to explanation by its ϵ-machine "theory", this simplest theory is typically larger, informationally speaking (C_μ), than the predictability benefit it confers. That said, the ϵ-machine is the best (simplest) theory. Thus, we use C_μ to define our notion of classical simplicity. It provides an interpretable ordering of processes—process A is simpler than process B when $C_\mu^A < C_\mu^B$.

We may also consider the recently proposed quantum-machine representation of processes [10–12]. The quantum-machine consists of a set $\{|\eta_k(L)\rangle\}$ of pure *signal states* that are in one-to-one correspondence with the classical causal states $\sigma_k \in \boldsymbol{\mathcal{S}}$. Each signal state $|\eta_k(L)\rangle$ encodes the set of length-L words that may follow σ_k, as well as each corresponding conditional probability. Fixing L, we construct quantum states:

$$|\eta_j(L)\rangle \equiv \sum_{w^L\in\mathcal{A}^L}\sum_{\sigma_k\in\boldsymbol{\mathcal{S}}} \sqrt{\Pr(w^L,\sigma_k|\sigma_j)}\ |w^L\rangle\,|\sigma_k\rangle \ , \qquad (2)$$

where w^L denotes a length-L word and $\Pr(w^L,\sigma_k|\sigma_j) = \Pr(X_{0:L} = w^L, \mathcal{S}_L = \sigma_k|\mathcal{S}_0 = \sigma_j)$. The resulting Hilbert space is the product $\mathcal{H}_w \otimes \mathcal{H}_\sigma$. Factor space $\mathcal{H}_\sigma$ is of size $|\boldsymbol{\mathcal{S}}|$, the number of classical causal states, with basis elements $|\sigma_k\rangle$. Factor space $\mathcal{H}_w$ is of size $|\mathcal{A}|^L$, with basis elements $|w^L\rangle = |x_0\rangle\cdots|x_{L-1}\rangle$.

The quantum measure of memory is the von Neumann entropy of the stationary state:

$$C_q = -\operatorname{Tr}(\rho\log\rho) \ , \qquad (3)$$

where $\rho = \sum_i \pi_i\,|\eta_i\rangle\,\langle\eta_i|$. This quantum analog of memory is generically less than the classical: $C_q \le C_\mu$. Also, due to the Holevo bound [10, 24], $\mathbf{E} \le C_q$. Though rare in process space, the classical and quantum informational sizes are equal exactly when both models are "maximally simple" : $\mathbf{E} = C_q = C_\mu$.

Ising Chain Simplicity The Ising spin-chain Hamiltonian is given by:

$$H = -\sum_{<i,j>}(Js_is_j + bs_i) \ , \qquad (4)$$

where s_i, the spin at site i, takes values $\{-1,+1\}$, J is the nearest-neighbor spin coupling constant, and b is the strength of the external magnetic field.

In equilibrium the bi-infinite chain of spin random variables defines a stationary stochastic process which has been analyzed using computational mechanics [25]. Importantly, spins obey a conditional independence: $\Pr(X_{0:\infty}|x_{-\infty:0}) = \Pr(X_{0:\infty}|x_0)$. That is, the "future" spins (right half) depend not on the entire past (left half) but only on the most recent spin x_0. The conclusion (see Supp. Materials) is that the two-state Markov chain process is minimally represented by the ϵ-machine in Fig. 1. Using Eq. (1), the statistical complexity is directly calculated as a function of p and q. Figure 2 shows that C_μ is a monotonically increasing function of temperature T: $1 - C_\mu \propto T^{-2}$ at high T. In particular, for the three processes chosen at temperatures $T_\alpha < T_\gamma < T_\delta$,

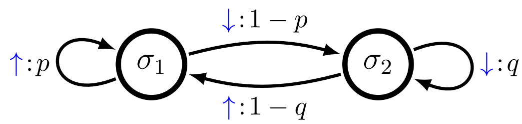

FIG. 1. The ϵ-machine for the nearest-neighbor Ising spin chain has two causal states σ_1 and σ_2. If the last observed spin x_0 is up ($s_0 = +1$) the current state is σ_1 and if it's down ($s_0 = -1$) is σ_2. If the current state is σ_1, with probability p the next spin observed is up and, if the current state is σ_2, with probability q the next spin observed is down.

$C_\mu^\alpha < C_\mu^\gamma < C_\mu^\delta$.

Consider now the quantum representation of these spin configurations. Each causal state is mapped to a pure quantum state that resides in a spin one-half space [13]:

$$
\begin{aligned}
|\sigma_1\rangle &= \sqrt{p}\,|\uparrow\rangle + \sqrt{1-p}\,|\downarrow\rangle \\
|\sigma_2\rangle &= \sqrt{1-q}\,|\uparrow\rangle + \sqrt{q}\,|\downarrow\rangle \ .
\end{aligned} \tag{5}
$$

(We use a more compact spin up/down notation, rather than the quantum machine notation of Eq. (2).) Intuitively, the quantum overlap accounts for the fact that the conditional predictions $\Pr(X_{0:\infty}|\sigma_1)$ and $\Pr(X_{0:\infty}|\sigma_2)$ share some subset of future outcomes. The density matrix is then:

$$
\rho = \pi_1 \, |\sigma_1\rangle \langle\sigma_1| + \pi_2 \, |\sigma_2\rangle \langle\sigma_2| \ . \tag{6}
$$

Computing the quantum analog $C_q = -\operatorname{Tr}(\rho \log \rho)$ as a function of temperature, Fig. 2 shows that this quantum size is generically well below the classical size C_μ. Thus, the quantum theory for the Ising chain is simpler than the classical: $C_q^\alpha < C_\mu^\alpha$, $C_q^\gamma < C_\mu^\gamma$, and $C_q^\delta < C_\mu^\delta$. Given the nature of progress in quantum information and computation [26, 27], it is notable, but perhaps no longer so surprising, that there exists such a quantum representational advantage.

Ambiguity of Simplicity Absolute sizes aside, what can we say about the associated process *rankings*? How does the notion of "simpler" survive the transition from classical to quantum description?

Observe (Fig. 2) that, unlike the classical measure C_μ, the quantum simplicity C_q is not monotonic in temperature: $C_q^\alpha < C_q^\delta < C_q^\gamma$. Moreover, the maximum C_q occurs at temperature $T_{C_q} \simeq 1.63$ while the excess entropy is maximized at temperature $T_{\mathbf{E}} \simeq 1.53$. These straightforward observations provide the kernel of several counterintuitive consequences.

First, what is the consequence of nonmonotonicity? Take the processes α and γ in Fig. 2. Classically and quantally, α is simpler than γ. In contrast, the ranking of processes γ and δ changes, $C_\mu^\gamma < C_\mu^\delta$ and $C_q^\gamma > C_q^\delta$.

In this way, even the familiar 1D Ising spin chain illustrates what is a general phenomenon—the ambiguity of

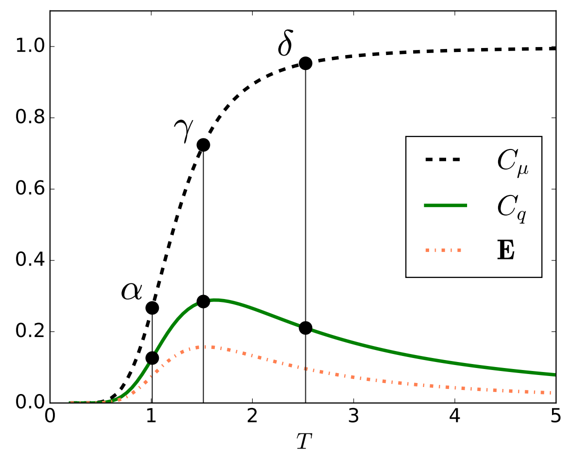

FIG. 2. Classical and quantum measures of Ising chain simplicity: Statistical complexity C_μ, quantum state complexity C_q, and excess entropy $\mathbf{E}$ versus temperature T in units of J/k_B at $b = 0.3$ and $J = 1$. ($C_\mu(T)$ and $\mathbf{E}(T)$ after Ref. [28] and $C_q(T)$ after Ref. [13].) Three particular spin processes are highlighted α, γ, and δ at temperatures T_α, T_γ, and T_δ.

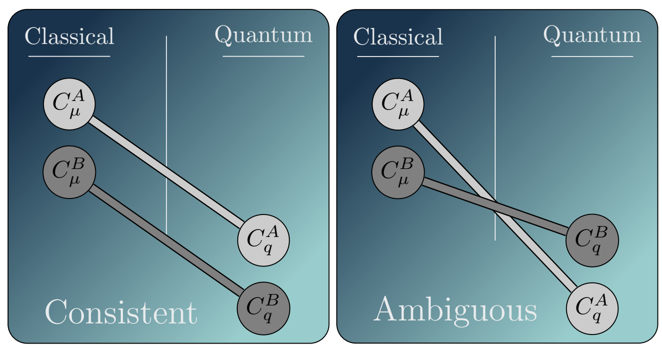

FIG. 3. (left) Classical and quantum rankings provide a consistent interpretation of which process is simpler. (right) Rankings reverse. And so, the question of simplicity is ambiguous.

simplicity. How general? Consider two generic processes A and B, for which no change in ranking occurs under the quantum lens. This indicates a *consistency* between the two representational viewpoints, at least with respect to processes A and B: $C_\mu^A > C_\mu^B$ and $C_q^A > C_q^B$. Figure 3(left) illustrates this circumstance. It can also be the case that the simplicity ranking of A and B *changes* when moving from classical to quantum representation. We refer to this as *ambiguity*. See Fig. 3(right). One concludes that the basic question—"Which process is simpler?"—no longer has a well defined answer.

How generic are consistency and ambiguity in the Ising spin chain parameter space? In Fig. 4 we construct an ambiguity diagram that compares all pairs of processes at temperatures T_1 and T_2 in the range $[0, 5]$. There, we fix the magnetic field $b = 0.3$ and coupling constant

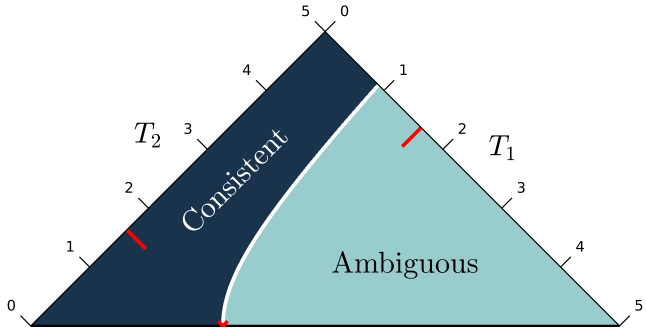

FIG. 4. Ambiguity diagram for Ising spin chain: Each point corresponds to a pair of Ising spin chains at temperatures T_1 and T_2 with $J = 1$ and $b = 0.3$. Given the inherent symmetry, the figure shows only half of the $T_1 \times T_2$ square. Consistency is found near the ($T = 0$) axes, while ambiguity dominates the remainder of parameter space. Curved boundary between these two regions ends at a temperature corresponding to $\max(C_q)$: $T_{C_q} \simeq 1.63$ (marked as a red dash).

$J = 1$. We find that the only consistent pairs are those within a shrinking envelope around the axes ($T_1 = 0$ and $T_2 = 0$). The bulk of parameter space, then, contains ambiguously ranked pairs. The singular feature of the diagram is the leftmost point along the boundary between the two regimes. This occurs at the temperature $T_{C_q} \simeq 1.63$ where we find the maximum value of C_q. Monotonicity of C_μ ensures that the transition between consistency and ambiguity depends on a reordering of C_q (not C_μ) values.

A notable case occurs when $b = 0$: there is no external-field induced symmetry-breaking. As a consequence, $C_\mu = 1$ for all temperatures and C_q is a decreasing function of temperature. This means that, for every pair of temperatures, we are at the border line of ambiguity. Classically they are as simple as each other, but quantally the system at the higher temperature is simpler.

Robustness of ambiguity One may object that this ambiguity is merely an artifact of the particular quantum construction or its size measure C_q. This is a valid concern, especially since minimality of this (or any other quantum representation) has not been established. Critically, the essence of ambiguity does not depend on this contingency, as we now show.

Denote by $\widetilde{C_q}$ the memory measure of an optimal quantum model[1] $\widetilde{Q}$ built according to some hypothetical, quantum scheme. Since $\widetilde{C_q}$, like C_q, is also bounded between $\mathbf{E}$ and C_μ [10, 24], we can define sufficient criteria for consistency and ambiguity between $\widetilde{C_q}$ and C_μ. We assume that the hypothetical model $\widetilde{Q}$ is no less efficient than the original quantum-machine: $\widetilde{C_q} \leq C_q$.

[1] The quantum model with minimum Von Neumann entropy over it's states.

Assume that for processes A and B, B is classically simpler. Then, the stronger criterion $\mathbf{E}^A > C_q^B$ ensures that any $\widetilde{Q}$ must yield consistency in rankings and is therefore, what we call, *certainly consistent*. See Fig. 5(left). Similarly, if $\mathbf{E}^B > C_q^A$, we know that any $\widetilde{Q}$ must yield an ambiguous ordering and is *certainly ambiguous*. See Fig. 5(right).

Figure 6 illustrates these stricter relations within the same Ising parameter region; compare Fig. 4. The central region does not satisfy either strict constraint. As expected, the certainly consistent (ambiguous) area is a proper subset of the consistent (ambiguous) area.

One concludes that no matter what future improvements may be found in quantum representations, these "certain" subregions are robust. This is a strong statement about how one can or cannot systematically rank the simplicity of systems classically and quantally. Again, the basic Ising spin chain is sufficiently rich to illustrate these new phenomena.

Discussion How common is ambiguity? The Supplementary Materials show that it is quite common in the analogous (nearest-neighbor, ferromagnetic) two-dimensional Ising system. Perhaps, however, the ambiguity of simplicity is special to spin systems. The Supplementary Materials establish that it is, in fact, a much more general phenomenon, by introducing a set of easily satisfied conditions such that two simplicity functions over a set of structured objects must yield ambiguous ordering. In particular, taking the space of all ϵ-machines as a set and C_μ and $\widetilde{C_q}$ as the two measures , we find that these conditions are satisfied. The general consequence is that either the two measures selected are trivially equal or ambiguity must exist. In other words, if the world is not ambiguous, quantum mechanics cannot simplify its explanation. One concludes that ambiguity is necessary for quantum simplification.

Closing Remarks The comparison of classical and quantum descriptions calls into question basic scientific practices that rely on a belief in the simplicity of the physical world. These two worlds disagree on simplicity ranking. Monitoring model simplicity is far from being the sole domain of physics. It is key in a variety of statistical inference tasks, notably in model selection [29]. Thus, the ambiguity of simplicity will have major practical consequences in a future that relies on quantum computing instead of classical.

Imagine competing models of some finite data $\mathcal{D}$. In Bayesian inference, one widely employed methodology, choosing one model over another requires specifying a prior probability distribution over models [30]. Such priors are commonly constructed to favor simpler models. Indeed, there is a long history of methods that avoid overfitting to data by incorporating simplicity measures into model selection, including Akaike's Information Criterion [31], Boltzmann Information Criterion [32], Min-

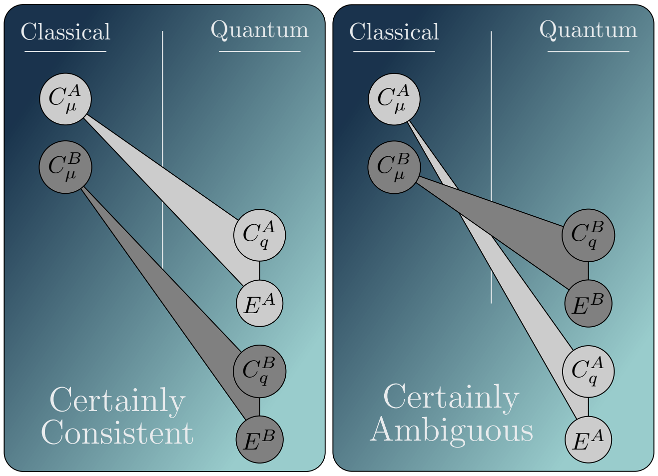

FIG. 5. Constraining hypothetical, as-yet-unknown frameworks for building quantum models $\widetilde{Q}$: Appealing to size measures C_q and $\mathbf{E}$ and without knowing any further details about $\widetilde{Q}$, we can still identify processes for which classical and quantum simplicity orderings must *certainly* be consistent or ambiguous. Cases exist that fall into neither of these stricter categories.

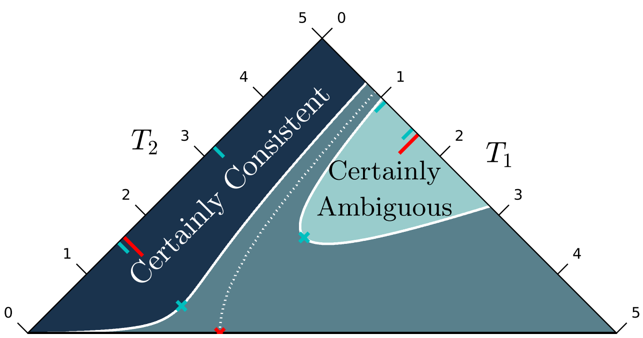

FIG. 6. Certain ambiguity diagram: Each point corresponds to a pair of Ising spin chains at temperatures T_1 and T_2 with $J = 1$ and $b = 0.3$. Dashed line marks Fig. 4's consitent-ambiguous border. Certainly consistent (ambiguous) is a proper subset of consistent (ambiguous). Local extrema of $\max(\mathbf{E})$ and $C_q = \max(\mathbf{E})$ along the new boundaries are marked with short blue lines at the corresponding temperatures. Long red lines mark the same values as in Fig. 4.

imum Description Length [33], and Minimum Message Length [34].

Classically, we may find that model A is simpler than B and increase its prior accordingly. Given that the two likelihoods $\Pr(\mathcal{D}|A)$ and $\Pr(\mathcal{D}|B)$ are similar enough, our inference identifies A as preferred. As we showed, the tables may turn dramatically when evaluating quantum models; we might find there that B is much simpler. We must then reconcile the fact that the quantum lens reveals a different answer.

We introduced the ambiguity of simplicity focusing on classical and quantum descriptions of classical processes. Quantum supremacy [2] suggests we go further to explore how (and if) ambiguity manifests when modeling quantum processes. This can be probed in the 1D quantum Heisenberg spin chain [35], for example. Measuring each spin in the z-direction yields a stochastic process—one that can be described classically or quantally. The Heisenberg spin chain is realized experimentally in the quasi-1D magnetic order found in antiferromagnetic $KCuF_3$ crystals [36–38]. One can then adapt the methods of 1D chaotic crystallography [39] to extract the ϵ-machine and quantum-machine descriptions of the quantum crystalline structure from scattering measurements. These and perhaps other experiments will provide an entrée to analyzing the ambiguity of simplicity in quantum systems.

Supplementary Materials: Details on spin system calculations and proof of the robustness theorem for the ambiguity of simplicity.

ACKNOWLEDGMENTS

We thank A. Aghamohammadi, M. Anvari, R. D'Souza, R. James, and D. Varn for useful conversations. This material is based upon work supported by, or in part by, the John Templeton Foundation under grant 52095 and the U. S. Army Research Laboratory and the U. S. Army Research Office under contracts W911NF-13-1-0390 and W911NF-13-1-0340.

[1] Sir Isaac Newton. *The Mathematical Principles of Natural Philosophy.* H. D. Symonds, 1729. A. Motte, trans.
[2] J. Preskill. arXiv:1203.5813.
[3] I. Boettcher, L. Bayha, D. Kedar, P. A. Murthy, M. Neidig, M. G. Ries, A. N. Wenz, G. Zurn, S. Jochim, and T. Enss. *Phys. Rev. Lett*, 116:045303, 2016.
[4] K. Fenech, P. Dyke, T. Peppler, M. G. Lingham, S. Hoinka, H. Hu, , and C. J. Vale. *Phys. Rev. Lett.*, 116:045302, 2016.
[5] J. P. Crutchfield and K. Young. *Phys. Rev. Let.*, 63:105–108, 1989.
[6] D. P. Feldman, C. S. McTague, and J. P. Crutchfield. *CHAOS*, 18(4):043106, 2008.
[7] P. W. Shor. *Algorithms for quantum computation: Discrete logarithms and factoring.* IEEE, 1994.
[8] D. Deutsch and R. Jozsa. Rapid solution of problems by quantum computation. *Proc. Roy. Soc. Ser. A*, 439:553, 1992.
[9] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. D. Sarma. Non-abelian anyons and topological quantum computation. *Rev. Mod. Physics*, 80(3):1083, 2008.
[10] M. Gu, K. Wiesner, E. Rieper, and V. Vedral. *Nature Comm.*, 3(762):1–5, 2012.
[11] J. R. Mahoney, C. Aghamohammadi, and J. P. Crutchfield. Occam's quantum strop: Synchronizing and com-

pressing classical cryptic processes via a quantum channel. *Scientific Reports*, 6, 2016.
[12] P. M. Riechers, J. R. Mahoney, C. Aghamohammadi, and J. P. Crutchfield. Minimized state complexity of quantum-encoded cryptic processes. *Phys. Rev. A*, 93(5):052317, 2016.
[13] W. Y. Suen, J. Thompson, A. J. Garner, V. Vedral, and M. Gu. arXiv:1511.05738.
[14] A. Monras, A. Beige, and K. Wiesner. arXiv:1002.2337.
[15] C. Aghamohammadi, J. R. Mahoney, and J. P Crutchfield. arXiv:1609.03650.
[16] A. J. P. Garner, Q. Liu, J. Thompson, V. Vedral, and M. Gu. arXiv:1609.04408.
[17] M. S. Palsson, M. Gu, J. Ho, H. M. Wiseman, and G. J. Pryde. arXiv:1602.05683.
[18] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* Wiley-Interscience, New York, second edition, 2006.
[19] A. N. Kolmogorov. *Dokl. Akad. Nauk. SSSR*, 119:861, 1958. (Russian) Math. Rev. vol. 21, no. 2035a.
[20] J. P. Crutchfield and D. P. Feldman. *CHAOS*, 13(1):25–54, 2003.
[21] J. P. Crutchfield. *Nature Physics*, 8(January):17–24, 2012.
[22] N. F. Travers and J. P Crutchfield. *J. Stat. Physics*, 145(5):1181–1201
[23] R. G. James, J. R. Mahoney, C. J. Ellison, and J. P. Crutchfield. *Phys. Rev. E*, 89:042135, Apr 2014.
[24] A. S. Holevo. *Problems Inform. Trans.*, 9:177–183, 1973.
[25] D. P. Feldman and J. P. Crutchfield. Discovering noncritical organization: Statistical mechanical, information theoretic, and computational views of patterns in simple one-dimensional spin systems. 1998. Santa Fe Institute Working Paper 98-04-026.
[26] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, 2010.
[27] J. Preskill. Lecture notes for Physics 229: Quantum information and computation. *California Institute of Technology*, page 16, 1998.
[28] D. P. Feldman. *Computational Mechanics of Classical Spin Systems.* PhD thesis, University of California, Davis, 1998. Published by University Microfilms Intl, Ann Arbor, Michigan.
[29] K. P. Burnham and D. R. Anderson. *Model selection and multimodel inference: A practical information-theoretic approach.* Springer Science & Business Media, 2003.
[30] C. C. Strelioff and J. P Crutchfield. *Phys. Rev. E*, 89(4):042119, 2014.
[31] H. Akaike. *IEEE Trans. Auto. Control*, 19(6):716–723, 1974.
[32] G. Schwarz. *Ann. Stat.*, 6(2):461–464, 1978.
[33] J. Rissanen. *Stochastic Complexity in Statistical Inquiry*, volume 15. World Scientific, Singapore, 1998.
[34] C. S. Wallace and F. P. Freeman. *J. R. Statist. Soc. B*, 49:240, 1987.
[35] R. J. Baxter. *Exactly Solved Models in Statistical Mechanics.* Academic Press, 1989.
[36] D. A. Tennant, R. A. Cowley, S. E. Nagler, and A. M. Tsvelik. *Phys. Rev. B*, 52:13368–13380, 1995.
[37] D. A. Tennant, S. E. Nagler, D. Welz, G. Shirane, and K. Yamada. *Phys. Rev. B*, 52:13381–13389, 1995.
[38] J. M. Maillet. *Seminaire Poincare*, X:139–177, 2007.
[39] D. P. Varn and J. P. Crutchfield. *Curr. Opin. Chem. Eng.*, 7:47–56, 2015.

Supplementary Materials
The Ambiguity of Simplicity
Cina Aghamohammadi, John R. Mahoney, and James P. Crutchfield

To ground the notion of simplicity, the main text couched the discussion in terms as physical (and familiar) as possible by considering the Ising spin chain from statistical physics [S1]—a model that historically played a critical role in understanding phase transitions [S2], spin glasses [S3], and lattice gasses [S4]. Its impact has reached well beyond physics, too, to ecology [S5], financial economics [S6], and neuroscience [S7]. Specifically, the main text focused on the one-dimensional nearest-neighbor Ising spin chain in the thermodynamic limit, showing how it inherently contains an ambiguous simplicity ordering. Here, we provide additional details underlying that analysis, generalize the result to show that ambiguity also appears in the, perhaps even more familiar, 2D Ising lattice, and finally establish the robustness of ambiguity via a theorem that lays out its most basic conditions.

ON SPIN CHAIN SIMPLICITY

Importantly, spins in the 1D chain obey a conditional independence: $\Pr(X_{0:\infty}|x_{-\infty:0}) = \Pr(X_{0:\infty}|x_0)$. That is, the "future" spins (right half) depend not on the entire past (left half) but only on the most recent spin x_0. Therefore, spin configurations resulting from the Hamiltonian in Eq. (4) can be modeled by a simple two-state Markov chain consisting of up (↑) and down (↓) states with self-transition probabilities [S8]:

$$p \equiv \Pr(\uparrow \mid \uparrow) = N_+/D \text{ and}$$
$$q \equiv \Pr(\downarrow \mid \downarrow) = N_-/D \ ,$$

where $N_\pm = \exp \beta(J \pm b)$ and:

$$D = \exp(\beta J) \cosh(\beta b) + \sqrt{\exp(-2\beta J) + \exp(2\beta J) \sinh(\beta b)^2} \ ,$$

with $\beta = 1/(k_B T)$.

Calculating the ϵ-machine via the causal-state equivalence relation is straightforward. There are exactly two causal states; except when $p = 1 - q$ where we find only one causal state. The conclusion is that the two-state Markov chain process is minimally represented by the ϵ-machine in Fig. 1. Using Eq. (1), the statistical complexity is directly calculated as a function of p and q:

$$C_\mu = -\Big(\frac{1-q}{2-p-q}\Big) \log_2 \Big(\frac{1-q}{2-p-q}\Big) - \Big(\frac{1-p}{2-p-q}\Big) \log_2 \Big(\frac{1-p}{2-p-q}\Big) \ .$$

Figure 2 showed that C_μ is a monotonically increasing function of temperature T: $1 - C_\mu \propto T^{-2}$ at high T. In particular, for the three processes chosen at temperatures $T_\alpha < T_\gamma < T_\delta$:

$$C_\mu^\alpha < C_\mu^\gamma < C_\mu^\delta \ .$$

Recall that Fig. 4 demonstrated that the presence of ambiguity is robust: There exist parameter regions in which the ambiguity is stable against alternative quantum representations—alternatives that arguably could lead to different simplicity metrics.

To address how common ambiguity is, consider ambiguity in the analogous (nearest-neighbor, ferromagnetic) two-dimensional Ising system. To answer this question we need to come back and look closely at the important difference between C_μ and C_q. While C_q is a smooth function of ϵ-machine's transition probabilities, this is not generally true for C_μ. Consider Fig. 1 for p and q close to $p = 1/2$. In this case, we have a uniform distribution over causal states and consequently $C_\mu \simeq 1$. For the quantum-machine, using Eq. (5), two states are very close to each other meaning $\langle \sigma_1 | \sigma_2 \rangle \simeq 1$. The consequence is $C_q \simeq 0$. Now, lets look at the case where $p = q = 1/2$. The two causal states are not distinguishable and we only have one causal state and as a result $C_\mu = 0$. We also have $|\sigma_1\rangle = |\sigma_2\rangle$ which leads to $C_q = 0$. The lesson here is that C_q smoothly tracks how distinguishable states are, but C_μ tracks if causal states are exactly distinguished or not.

Recall that for the 1D Ising chain at high temperatures $p \neq q$; in particular, though they are both are close to $1/2$, they are never equal. The consequence is $C_q \simeq 0$ and $C_\mu = 1$. What about for the 2D Ising model? At the extreme

$T = 0$ and for any nonzero value of external field, the ground state will be in uniform alignment with the field. This means that any random variable constructed from spin variables must have vanishing entropy. Lacking a complete computational mechanics of structure in two-dimensional patterns (though see Ref. [S9, S10]), it is still clear that any analog of statistical complexity (and thereby C_q) will vanish at $T = 0$ for such uniform configurations.

At very high T, though, spins become increasingly uncorrelated and the configurational conditional probability distribution associated with each causal state approaches uniformity, *but as a set the distributions remain distinct.* That is, for any sufficiently high finite temperature, the system has some, perhaps weak, correlation that keeps the distributions from becoming identical. In other words, causal states in this regime remain probabilistically distinct. So, as with the 1D case, at very high temperature ($T \gg 1$, but $T \neq \infty$) $C_\mu(T)$ is not zero.

What can we say about C_q in this limit? For high $T \gg 1$ spin randomness makes the quantum states $\{|\eta\rangle\}$ (Eq. (2)) more and more indistinguishable. And so, their increasing overlaps $\langle \eta_i | \eta_j \rangle \to 1$, driving C_q to zero monotonically. The conclusion is that for the 2D Ising, at $T \approx 0$ and $T \gg 1$, we have the same qualitative picture for the simplicity measures as depicted in Fig. 2. This brief argument says that ambiguity exists in the 2D Ising spin model as well.

AMBIGUITY ROBUSTNESS THEOREM

First, we lay bare the mathematical argument and then we interpret it in terms of the physical setting of the main text.

Consider a set of objects S and two functions over the set $F_1 : S \to G$ and $F_2 : S \to G$. Space G consists of elements that can be compared as follows.

If there exists $s_1, s_2 \in S$, such that $F_1(s_1) > F_1(s_2)$ and $F_2(s_1) < F_2(s_2)$, then we say these functions are *ambiguous* over S.

We define three conditions for the set and functions.

Condition A The two functions map onto the whole space G: $F_1(S) = G$ and $F_2(S) = G$.

Condition B For all $g \in G$ there exists $x \in S$ such that $F_1(x) = F_2(x) = g$.

Condition C Assume $\preceq$ is a dense, total order on space G.

Theorem 1. *Given two functions F_1 and F_2 that map set S to space G and satisfy Conditions A, B, and C: No ambiguity implies that for all $x \in S$, $F_1(x) = F_2(x)$.*

Proof. *We prove the contrapositive by contradiction. Assume there exists $x \in S$ such that $F_1(x) \neq F_2(x)$. Without loss of generality, let $F_1(x) \succeq F_2(x)$. Since $\preceq$ is a dense total order on G, there is $g \in G$ such that $F_1(x) \succeq g \succeq F_2(x)$. By Condition B, there exists $y \in S$ such that $F_1(y) = F_2(y) = g$. Trivially then, $F_1(x) \succeq F_1(y)$ and $F_2(x) \preceq F_2(y)$. This demonstrates ambiguity and completes the proof.*

We can interpret this in the setting of stationary processes with measures C_μ and $\widetilde{C_q}$ and discuss the space of all possible quantum sizes. More specifically, consider the case $F_1 = C_\mu$ and $F_2 = \widetilde{C_q}$. We know that for any value $y \in \mathbb{R}$, there exists an ϵ-machine with $C_\mu = \widetilde{C_q} = \mathbf{E} = y$. This satisfies the assumption. Then, our results say that if the world is not ambiguous, the two measures are equivalent. In other words, the quantum advantage $C_\mu / \widetilde{C_q}$ *requires* ambiguity.

[S1] S.. G. Brush. *Rev. Mod. Phys.*, 39(4):883, 1967.
[S2] R. Peierls. *Math. Proc. Cambridge Phil. Soc.*, 32(03):477–481, 1936.
[S3] S. F. Edwards and P. W. Anderson. *J. Physics F: Metal Physics*, 5(5):965, 1975.
[S4] T.-D. Lee and C.-N. Yang. *Phys. Rev.*, 87(3):410, 1952.
[S5] A. E. Noble, J. Machta, and A. Hastings. *Nature Comm.*, 6:6664, 2015.
[S6] D. Sornette. *Rep. Prog. Physics*, 77(6):062001, 2014.
[S7] E. Schneidman, M. J. Berry, R. Segev, and W. Bialek. *Nature*, 440(7087):1007–1012, 2006.
[S8] D. P. Feldman and J. P. Crutchfield. Discovering non-critical organization: Statistical mechanical, information theoretic, and computational views of patterns in simple one-dimensional spin systems. 1998. Santa Fe Institute Working Paper 98-04-026.
[S9] D. P. Feldman. *Computational Mechanics of Classical Spin Systems.* PhD thesis, University of California, Davis, 1998. Published by University Microfilms Intl, Ann Arbor, Michigan.
[S10] D. P. Feldman and J. P. Crutchfield. *Phys. Rev. E*, 67(5):051103, 2003.